X - 1

AB

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04Q 11/04 — A1

(11) International Publication Number: WO 99/23853

(43) International Publication Date: 14 May 1999 (14.05.99)

(21) International Application Number: PCT/US98/23535

(22) International Filing Date: 4 November 1998 (04.11.98)

(30) Priority Data:
08/965,055 5 November 1997 (05.11.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 08/965,055 (CIP)
Filed on 5 November 1997 (05.11.97)

(71) Applicant *(for all designated States except US)*: NOKIA IP INC. [US/US]; 232 E. Java Drive, Sunnyvale, CA 94089-1318 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DECANIO, Thomas, A. [US/US]; 5451 Alamo Terrace, Fremont, CA 94555 (US). LYON, Thomas [US/US]; 1400 Edgewood Drive, Palo Alto, CA 94301 (US). NEWMAN, Peter [US/US]; 750 N. Shoreline Boulevard #124, Mountain View, CA 94043 (US). MINSHALL, Greg [US/US]; 553 Tyndall, Los Altos, CA 94022 (US). HINDEN, Robert [US/US]; 3271 Murray Way, Palo Alto, CA 94303 (US). LIAW, Fong, Ching [US/US]; 1904 Messina Drive, San Jose, CA 95132 (US). HOFFMAN, Eric [US/US]; 740 Sixth Avenue, Redwood City, CA 94063 (US).

(74) Agents: DINH, Truong, T. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).

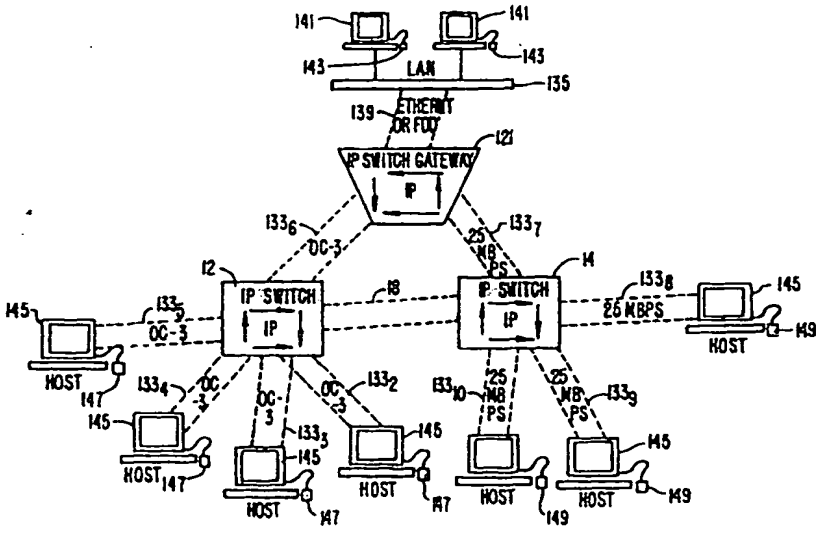
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MULTIPORT INTERFACES FOR A NETWORK USING INVERSE MULTIPLEXED IP SWITCHED FLOWS



BEST AVAILABLE COPY

(57) Abstract

A method and apparatus for dynamically shifting between switching and routing packets efficiently to provide high packet throughput. The present invention designates multiple ports of an IP switched router for communication as a single interface with a second adjacent IP switched router to provide increased bandwidth between the network interfaces on the IP switched routers. The multiple designated ports are monitored by both IP switched routers for communications from the other. Data flows are queued up and inverse multiplexed over the multiple ports to optimize the available bandwidth. The inverse multiplexing is done at layer 2 of the OSI reference model, so that layer 3 does not have to know about any reallocation among the multiple ports.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

MULTIPORT INTERFACES FOR A NETWORK USING INVERSE MULTIPLEXED IP SWITCHED FLOWS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority from commonly-owned U.S. provisional patent application no. 60/030,348 filed on November 6, 1996, the disclosure of which is herein incorporated by reference for all purposes. This application is also a continuation-in-part application of commonly-assigned U.S. patent application no. 08/792,183 filed on January 30, 1997 which is a continuation-in-part of U.S. patent no. 08/597,520 filed on January 31, 1996, the disclosures of which are herein incorporated by reference for all purposes.

COPYRIGHT NOTICE

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

BACKGROUND OF THE INVENTION

The present invention relates to the field of network communications. In particular, a specific embodiment of the invention relates to improving the bandwidth of communications between the network interfaces on two adjacent "IP switched routers."

Local area network (LAN) switches have been conventionally used as a quick, relatively inexpensive way to relieve congestion on shared-media LAN segments to more effectively manage traffic and allocate bandwidth within a LAN than shared-media hubs or simple bridges. LAN switches operate as datalink layer (layer 2 of the OSI reference model) packet-forwarding hardware engines, dealing with media access control (MAC) addresses and performing simple table look-up functions. Although switch-based networks are able to offer greater throughput, they continue to suffer from problems such

as broadcast flooding and poor security. Routers, which operate at the network-layer (layer 3 of the OSI reference model), are still required to solve these types of problems. However, fast switching technology is overwhelming the capabilities of current routers, creating router bottlenecks. The traditional IP packet-forwarding device on which the Internet is based, the IP router, is showing signs of inadequacy. In addition, routers are often expensive, complex, and of limited throughput, as compared to emerging switching technology. To support the increased traffic demand of large enterprise-wide networks and the Internet, IP routers need to operate faster and cost less. In current routers, multiple paths are available between two routers and a router may switch data between one or more paths by making a decision at layer 3 of the OSI reference model. These traditional routers achieve throughput in the hundreds of thousands of packets-per-second range. However, as the need for even greater throughput increases and advanced functionalities required by more types of traffic are enabled in IP, traditional IP routers will not suffice as packet-forwarding devices, especially since these routers are often limited by their processor-intensive designs.

From the above, it is seen that another approach for avoiding bottlenecks and increasing packet throughput between nodes is needed.

SUMMARY OF THE INVENTION

The present invention designates multiple ports of an IP switched router for communication as a single interface with a second adjacent IP switched router to provide increased bandwidth between the network interfaces on the IP switched routers. The multiple designated ports are monitored by both IP switched routers for communications from the other. Data flows are queued up and inverse multiplexed over the multiple ports to optimize the available bandwidth. The inverse multiplexing is done at layer 2 of the OSI reference model, so that layer 3 does not have to know about any reallocation among the multiple ports.

According to an embodiment, the present invention provides a method for transmitting packets over a multiport interface between an upstream node and a downstream node in a network, where the downstream node is downstream from the upstream node. The method includes the steps of establishing a multiport interface that includes multiple sub-ports between the upstream node and the downstream node, receiving a packet at the downstream node, and performing a flow classification at the

downstream node on the packet to determine whether the packet belongs to a specified flow that should be redirected in the upstream node to the multiport interface. The method also includes the steps of selecting a free label for one of the multiple sub-ports at the downstream node, and informing the upstream node that future packets belonging to the specified flow should be sent with the selected free label attached.

According to another embodiment, the present invention provides a computer program product that enables dynamic shifting between routing and switching in a network having an upstream node and a downstream node. The downstream node is downstream from the upstream node. The computer program product includes computer-readable code that establishes a multiport interface which includes multiple sub-ports between the upstream node and the downstream node, and computer-readable code that performs a flow classification on a packet at the downstream node to determine whether the packet belongs to a specified flow that should be redirected in the upstream node to the multiport interface. The computer program product also includes computer-readable code that selects a free label for one of the multiple sub-ports at the downstream node, computer-readable code that informs the upstream node that future packets belonging to the specified flow should be sent with the selected first free label attached, and a tangible medium that stores the computer-readable codes.

These and other embodiments of the present invention, as well as its advantages and features, are described in more detail in conjunction with the text below and attached figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram illustrating multiple port connections between two "IP switched routers," in accordance with a specific embodiment of the present invention;

Fig. 2 is a diagram illustrating the queuing of multiple flows for the multiple ports, in accordance with the specific embodiment of the present invention;

Fig. 3 illustrates one of the many network configurations possible in accordance with the present invention;

Fig. 4a is a system block diagram of a typical computer system 151 that may be used as switch controller 12a in basic switching unit 12 (as shown in Fig. 1) to execute a specific embodiment of the system software of the present invention;

Fig. 4b is a general block diagram of an architecture of an ATM switch 3

(the example shows a 16-port switch) that may be used as the switching hardware engine of a basic switching unit according to an embodiment of the present invention;

Fig. 5a is a simplified diagram generally illustrating the initialization procedure in each system node according to an embodiment of the present invention;

Fig. 5b is a simplified diagram that generally illustrates the operation of a system node according to an embodiment of the present invention;

Fig. 6a is a diagram generally illustrating the steps involved in labelling a flow in a system node according to an embodiment of the present invention;

Fig. 6b is a diagram generally illustrating the steps involved in switching a flow in a basic switching unit according to an embodiment of the present invention;

Fig. 6c is a diagram generally illustrating the steps involved in forwarding a packet in a system node according to an embodiment of the present invention;

Fig. 7a is a simplified diagram generally illustrating the multiport interface establishing procedure in a basic switching unit according to an embodiment of the present invention;

Fig. 7b is a diagram generally illustrating some of the steps involved in determining whether a flow should be switched in basic switching units according to an embodiment of the present invention;

Fig. 7c is a diagram generally illustrating the steps involved in labelling a flow in the upstream link for a designated multiport interface in a basic switching unit, such as shown by label flow step 366 of Fig. 7b according to an embodiment of the present invention; and

Fig. 7d is a simplified diagram that generally illustrates some of the steps of the operation of the basic switching unit according to the specific embodiment of the present invention.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

CONTENTS

I. General
A. Inverse Multiplexed Transmission of Flow Labelled Packets
B. Flow Classification, IFMP, GSMP in the Specific Embodiment
II. System Hardware
A. Controller Hardware
B. Switching Hardware
III. System Software Functionality
A. Configuration of Multiport Interfaces
B. Flow Distribution on Multiport Interfaces
IV. Conclusion

I. General

The present invention provides for a multiport interface made of two or more sub-ports used as a single interface, with flow-by-flow inverse multiplexing, to provide at layer 2 of the OSI reference model very high speed trunking capability between "IP switched routers," also referred to as "basic switching units." The multiport interface increases the effective bandwidth for transmitting packets in a network. The method and apparatus will find particular utility and is illustrated herein as it is applied in the high throughput flow-based transmission of IP packets capable of carrying voice, video, and data signals over a local area network (LAN), metropolitan area networks (MAN), wide area network (WAN), Internet, or the like, but the invention is not so limited. The invention will find use in a wide variety of applications where it is desired to transmit packets over a network.

A. Inverse Multiplexed Transmission of Flow Labelled Packets

In accordance with specific embodiments of the present invention, Fig. 1 illustrates two switch controllers 12a and 14a, respectively coupled to a switching engine 12b and 14b. Each corresponding pair of switch controller and switching engine form what is referred to as an "IP switched router" or a "basic switching unit" (basic switching unit 12 includes switch controller 12a and switching engine 12b, and basic switching unit 14 includes switch controller 14a and switching engine 14b), a specific embodiment of

which is described in further detail below. Although referred to as a "switching" unit, it should be recognized that the basic switching unit of the system via system software installed on its switch controller dynamically provides both layer 2 switching functionality as well as layer 3 routing and packet forwarding functionality. In this particular example, the switching engine, which utilizes conventional and currently available asynchronous transfer mode (ATM) switching hardware, is an ATM switch. The ATM switching hardware providing the switching engine of the basic switching unit operates at the datalink layer (layer 2 of the OSI reference model). Any of the software normally associated with the ATM switch that is above the ATM Adaptation Layer type 5 (AAL-5) is completely removed. Thus, the signalling, any existing routing protocol, and any LAN emulation server or address resolution servers, etc. are removed. Of course, other switching technologies such as for example fast packet switching, frame relay, 100BaseT Fast Ethernet, Gigabit Ethernet, Fiber Distributed Data Interface (FDDI) or others may be used to provide the switching engine of the basic switching unit, depending on the application. The switch controller is a computer having multiple network adapter or network interface cards (NICs) connected to the switching engine via multiport interface 18. System software is installed in basic switching unit, more particularly in the computer serving as switch controller. The switching engine serves to perform high-speed switching functions when required by the basic switching unit, as determined by the system software. The switching capability of the switching system is limited only by the hardware used in the switching engine. Accordingly, the present embodiment of the invention is able to take advantage of the high-speed, high capacity, high bandwidth capabilities of ATM technology. In addition to performing standard connectionless IP routing functions at layer 3, the switch controller also makes flow classification decisions for packets on a local basis, as described generally below.

In accordance with a specific embodiment of the present invention, each of these basic switching units can also communicate with other nodes in a network or other networks or servers via ports 16, for example. In one possible network configuration, a trunk interface between the two basic switching units may be used. The switching engine of each basic switching unit has multiple physical ports, each being capable of being connected to a variety of devices, including for example data terminal equipment (DTE), data communication equipment (DCE), servers, switches, gateways, etc. One of these multiple ports, for example port 1, is used to provide the communication link between the

switch controller and the switching engine. Two or more of these multiple ports may be used as a trunk interface to form a single multiport interface 18. For example, ports 8, 9, 10 and 12 of the basic switching unit 12 may be designated to be sub-ports of multiport interface 18, as described further below. Multiport interface 18 is created by combining several ports of the switching engine of a basic switching unit and having it appear as a single interface to the switch controller of the basic switching unit.

Fig. 2 illustrates an example of flow-by-flow inverse multiplexing across the multiple sub-ports of multiport interface 18. By way of illustration, flows 1-12 are shown being allocated to different queues for ports 8, 9, 10 and 12 of switching engine 12b. More specifically in this example, flows 4, 5 and 8 have been allocated to sub-port 8; flows 3, 9 and 11 have been allocated to sub-port 9; flows 2 and 7 have been allocated to sub-port 10; and flows 1, 6, 10 and 12 have been allocated to sub-port 12. The bandwidth can be maximized by evenly spreading the flows across the multiple sub-ports to the extent possible.

In addition, there is an optimization between (a) the multiport interface or trunk 18 between switch controllers 12 and 14, and (b) the multiple connections to other nodes. If the bandwidth required on other ports 16 to other nodes increases, one of the ports allocated to multiport interface 18 could be reallocated to communicate with another node. Conversely, if the bandwidth of traffic to other nodes decreases, more ports could be allocated to the multiport trunk interface 18 between switch controllers 12 and 14.

With specific embodiments of the present invention, various network configurations may be implemented to provide end-to-end seamless IP traffic flow, with the network configurations featuring high bandwidth and high throughput between network interfaces on basic switching units 12 and 14 via the flow-by-flow inverse multiplexing over multiport interface 18 established between basic switching units 12 and 14. For example, Fig. 3 illustrates one of the many network configurations possible in accordance with the present invention. Of course, many alternate configurations are possible. In one embodiment, multiport interface 18 could be used between basic switching units 12 and 14 as shown in Fig. 3, according to the present invention. Basic switching units, switch gateway units, and system software allow users to build flexible IP network topologies targeted at the workgroup, campus, and WAN environments for high performance, scaleable solution to current campus backbone congestion problems.

More specifically, Fig. 3 shows a simplified diagram of a high performance

workgroup environment in which several host computers 145 are connected via ATM links 133_m to multiple basic switching units 12 and 14, which each connect to a switch gateway unit 121 that connects to a LAN 135 with user devices 141. In this configuration, a first basic switching unit 12 connects to a second basic switching unit 14 via multiport interface 18, such as seen in Fig. 1. Multiple host computers 145 connect to the first basic switching unit 12 via respective 155 Mbps ATM links 133_x (where $x = 2$ to 5) through respective ATM NICs 147. In addition, multiple host computers 145 connect to the second basic switching unit 14 via respective 25 Mbps ATM links 133_y (where $y = 8$ to 10) through respective ATM NICs 149. As discussed above, host computers 145 equipped with ATM NICs are installed with a subset of the system software, enabling the TCP/IP hosts to connect directly to a basic switching unit. The first and second basic switching units 12 and 14 connect to switch gateway unit 121 via ATM links 133_6 (155 Mbps) and 133_7 (25 Mbps) respectively. Connection of the first and second basic switching units 12 and 14 to switch gateway unit 121 via an Ethernet (e.g., 10BaseT) or FDDI link 139 enables users of host computers 145 to communicate with users devices 141 attached to LAN 135. User devices 141 may be PCs, terminals, or workstations having appropriate NICs 143 to connect to any Ethernet or FDDI LAN 135. The workgroup of host computers is thereby seamlessly integrated with the rest of the campus network.

It is noted that a "switch gateway unit," which is similar to a basic switching unit without a switching engine, includes a gateway switch controller and IFMP software installed on the gateway switch controller, in accordance with a specific embodiment. The gateway switch controller includes multiple network adaptors or NICs, and an ATM NIC. Switch gateway unit serves as an access device to enable connection of existing LAN and backbone environments to a network of basic switching units. Accordingly, the NICs of the switch gateway unit may be of different types, such as for example Ethernet NICs, Fast Ethernet NICs, FDDI NICs, and others, or any combination of the preceding. Of course, the use of particular types of NICs depends on the types of existing LAN and backbone environments to which switch gateway unit provides access. It is recognized that multiple LANs may be connected to a switch gateway unit. The ATM NIC allows the switch gateway unit to connect via an ATM link to a basic switching unit. Of course, others of the other multiple NICs may also be ATM NICs to provide a connection from the switch gateway unit to another switch gateway. In addition to basic switching units and switch gateway units, networks utilizing the present invention also may also include high

performance host computers, workstations, or servers that are appropriately equipped. In particular, a subset of the IFMP software can be installed on a host computer, workstation, or server equipped with an appropriate ATM NIC to enable a host to connect directly to a basic switching unit.

According to specific embodiments of the present invention, system software on a switch controller of a basic switching unit can create and delete multiport interfaces and then direct the switching engine to switch flows on the multiport interface to implement the inverse multiplexing of flows over multiport interface 18 and adds complete IP routing functionality on top of ATM switching hardware instead of any existing conventional ATM switch control software, to control the ATM switch such that the flows are appropriately multiplexed over the sub-ports of interface 18. Therefore, the present system is capable of moving between network layer IP routing when needed and high throughput datalink layer flow switching over interface 18 when possible in order to create high speed and capacity packet transmission in an efficient manner without the problem of router bottlenecks. Using a multiport interface 18 between adjacent IP switched routers, the packet throughput between their attached network interfaces may reach millions of IP packets-per-second, which is an order of magnitude faster than with traditional IP routers.

B. Flow Classification, IFMP, GSMP in the Specific Embodiment

Using the Ipsilon Flow Management Protocol (IFMP), which is described in further detail in commonly-assigned U.S. patent application no. 08/597,520, a system node (such as a basic switching unit, switch gateway unit, or host computer/server/workstation) can classify IP packets as belonging to a "flow" of similar packets based on certain common characteristics. A flow is a sequence of packets sent from a particular source to a particular (unicast or multicast) destination that are related in terms of their routing and any local handling policy they may require. The present invention efficiently permits different types of flows to be handled differently, depending on the type of flow, and enables the inverse multiplexing of different flows over different sub-ports of multiport interface 18, depending on the bandwidth available on each designated sub-port. Some types of flows may be handled by mapping them into individual ATM connections using the ATM switching engine to perform high speed switching of the packets over multiport interface 18. Flows such as for example those carrying real-time traffic, those with quality of service requirements, or those likely to have a long holding time, may be configured to

be switched whenever possible. Other types of flows, such as for example short duration flows or database queries, may be handled by connectionless IP routing. A particular flow of packets may be associated with a particular ATM label (i.e., a virtual path identifier (VPI) and virtual channel identifier (VCI)). It is assumed that virtual channels are unidirectional so an ATM label of an incoming direction of each link is owned by the input port to which it is connected. Each direction of transmission on a link is treated separately. Of course, flows travelling in each direction are handled by the system separately but in a similar manner.

Flow classification is a local policy decision. When an IP packet is received by a system node, the system node transmits the IP packet via the default channel. The node also classifies the IP packet as belonging to a particular flow, and accordingly decides whether future packets belonging to the same flow should preferably be switched directly in the ATM switching engine or continue to be forwarded hop-by-hop by the router software in the node. If a decision to switch a flow of packets is made, the flow must first be labelled. To label a flow, the node selects for that flow an available label (VPI/VCI) of the input port on which the packet was received. The node which has made the decision to label the flow then stores the label, flow identifier, and a lifetime, and then sends an IFMP redirect message upstream to the previous node from which the packet came. The flow identifier contains the set of header fields that characterize the flow. The lifetime specifies the length of time for which the redirection is valid. Unless the flow state is refreshed, the association between the flow and label is deleted upon the expiration of the lifetime. Expiration of the lifetime before the flow state is refreshed results in further packets belonging to the flow to be transmitted on the default forwarding channel between the adjacent nodes. A flow state is refreshed by sending upstream a redirect message having the same label and flow identifier as the original and having another lifetime. The redirect message requests the upstream node to transmit all further packets that have matching characteristics to those identified in the flow identifier via the virtual channel specified by the label. The redirection decision is also a local decision handled by the upstream node, whereas the flow classification decision is a local decision handled by the downstream node. Accordingly, even if a downstream node requests redirection of a particular flow of packets, the upstream node may decide to accept or ignore the request for redirection. In addition, redirect messages are not acknowledged. Rather, the first packet arriving on the new virtual channel serves to indicate that the redirection request has been accepted.

Different encapsulations are used for the transmission of IP packets that belong to particular labelled flows on an ATM data link, depending on the different flow type of the flows. In the present example of Fig. 2, twelve types of encapsulations are used to transmit IP packets belonging to the twelve types of flows to be multiplexed over multiport interface 18.

In addition to using IFMP to classify and redirect flows, a system node such as a basic switching engine may utilize the General Switch Management Protocol (GSMP, also described in detail in U.S. patent application no. 08/597,520) to establish communication over the ATM link between the switch controller and ATM hardware switching engine of the basic switching unit and thereby enable flow-by-flow multiplexed layer 2 switching when possible and layer 3 IP routing and packet forwarding when necessary. In particular, GSMP is a general purpose, asymmetric protocol to control the switching engine, e.g., the ATM switch. That is, the switch controller acts as the master with the ATM switch as the slave. GSMP runs on a virtual channel established at initialization across the ATM link between the switch controller and the ATM switch. A single switch controller may use multiple instantiations of GSMP over separate virtual channels to control multiple ATM switches. Also included in GSMP is a GSMP adjacency protocol, which is used to synchronize state across the ATM link between the switch controller and the ATM switch, to discover the identity of the entity at the other end of the link, and to detect changes in the identity of that entity.

GSMP allows the switch controller to establish and release connections across the ATM switch, add and delete leaves on a point-to-multipoint connection, manage switch ports, request configuration information, and request statistics (such as the level of traffic on each port). GSMP also allows the ATM switch to inform the switch controller of events such as a link going down. In accordance with a specific embodiment of the present invention, a switch controller may use GSMP to configure multiport interface 18 and to direct the switching engine to switch flows on the multiport interface 18 such that the switching engine distributes the flows across the individual sub-ports designated by the configuration. Creation and deletion of multiport interfaces is done at the switch controller with configuration information being stored in, for example, non-volatile memory in the switch controller. The bandwidth and performance of the multiport interface approaches that of a single interface having a bandwidth equal to the sum of that of the individual sub-ports. For example, as seen in Fig. 2, combining four OC3 (155 Mbps) interfaces into a

multiport interface 18 creates an equivalent interface having a virtual bandwidth of an OC12 (622 Mbps) connection. The multiport interface 18 is useful in relieving network congestion that might result in trunk connections between basic switching units, such as when traffic from many LANs (e.g., Fast Ethernets) on one of the basic switching units might cross a trunk before reaching an important server or server farm.

Although shown in connection with distributing multiple flows as a preferred embodiment, the present invention can also be used to allocate bandwidth across multiple ports where other communication methods are used. In this situation, the advantage of being able to use multiple ports invisible to layer 3 of the OSI reference model provides advantages by reducing the overhead and time required to do the transmissions and reconfigure as necessary.

Switching engine 12b is assumed to contain multiple ports, where each physical port is a combination of an input port and an output port. ATM cells arrive at the ATM switch from an external communication link on incoming virtual channels at an input port, and depart from the ATM switch to an external communication link on outgoing virtual channels from an output port. As mentioned earlier, virtual channels on a port or link are referenced by their VPI/VCI. A virtual channel connection across an ATM switch is formed by connecting an incoming virtual channel (or root) to one or more outgoing virtual channels (or branches). Virtual channel connections are referenced by the input port on which they arrive and the VPI/VCI of their incoming virtual channel. In the switch, each port has a hardware look-up table indexed by the VPI/VCI of the incoming ATM cell, and entries in the tables are controlled by a local control processor in the switch.

II. System Hardware

A. Controller Hardware

Fig. 4a is a system block diagram of a typical computer system 151 that may be used as switch controller 12a in basic switching unit 12 (as shown in Fig. 1) to execute the system software of the present invention. Fig. 4a also illustrates an example of the computer system that may be used as switch gateway controller of switch gateway unit 121 (as shown in Fig. 3), as well as serving as an example of a typical computer which may be used as a host computer/server/workstation loaded with a subset of the IFMP software. Of course, it is recognized that other elements such as a monitor, screen, and

keyboard are added for the host.

As shown in Fig. 4a, computer system 151 includes subsystems such as a central processor 169, system memory 171, I/O controller 173, fixed disk 179, network interface 181, and read-only memory (ROM) 183. Of course, the computer system 151 optionally includes monitor 153, keyboard 159, display adapter 175, and removable disk 177, for the host. Arrows such as 185 represent the system bus architecture of computer system 151. However, these arrows are illustrative of any interconnection scheme serving to link the subsystems. For example, a local bus could be utilized to connect central processor 169 to system memory 171 and ROM 183. Configuration information for creation of multiport interfaces may be stored, for example, on ROM 183. Other computer systems suitable for use with the present invention may include additional or fewer subsystems. For example, another computer system could include more than one processor 169 (i.e., a multi-processor system) or a cache memory.

In an embodiment of the invention, the computer used as the switch controller can be a standard Intel-based central processing unit (CPU) machine equipped with a standard peripheral component interconnect (PCI) bus, as well as with an ATM network adapter or network interface card (NIC). The computer is connected to the ATM switch via a 155 Megabits per second (Mbps) ATM link using the ATM NIC. In this embodiment, the system software is installed on fixed disk 179 which is the hard drive of the computer. As recognized by those of ordinary skill in the art, the system software may be stored on a CD-ROM, floppy disk, tape, or other tangible media that stores computer-readable code.

Computer system 151 shown in Fig. 4a is but an example of a computer system suitable for use (as the switch controller of a basic switching unit, as the switch gateway controller of a switch gateway unit, or as a host computer/server/workstation) with the present invention. Other configurations of subsystems suitable for use with the present invention will be readily apparent to one of ordinary skill in the art. In addition, switch gateway unit may be equipped with multiple other NICs to enable connection to various types of LANs. Other NICs or alternative adaptors for different types of LAN backbones may be utilized in switch gateway unit. For example, SMC 10M/100M Ethernet NIC or FDDI NIC may be used.

Without in any way limiting the scope of the invention, Table 1 provides a list of commercially available components which are useful in operation of the controller,

according to the above embodiments. It will be apparent to those of skill in the art that the components listed in Table 1 are merely representative of those which may be used in association with the inventions herein and are provided for the purpose of facilitating assembly of a device in accordance with one particular embodiment of the invention. A wide variety of components readily known to those of skill in the art could readily be substituted or functionality could be combined or separated.

Table 1: Controller Components

Microprocessor	Intel Pentium 133 MHz processor
System memory	16Mbyte RAM/256K cache memory Motherboard Intel Endeavor motherboard
ATM NIC	Zeitnet PCI ATM NIC (155 Mbps)
Fixed or Hard disk Drives	500Mbyte IDE disk standard floppy, CD-ROM drive
Power supply	standard power supply
Chassis	standard chassis

B. Switching Hardware

As discussed above, the ATM switch hardware provides the switching engine 12b of basic switching unit 12, in accordance with a specific embodiment. The ATM switching engine utilizes vendor-independent ATM switching hardware. However, the ATM switching engine according to the present invention does not rely on any of its usual connection-oriented ATM routing and signaling software (SSCOP, Q.2931, UNI 3.0/3.1, and P-NNI). Rather, any ATM protocols and software are completely discarded, and the basic switching unit relies on the system software to create and delete multiport interfaces and to control the ATM switching engine for inverse multiplexing of flows. The system software is described in detail later.

Separately available ATM components may be assembled into a typical ATM switch architecture. For example, Fig. 4b is a general block diagram of an architecture of an ATM switch 12b (the example shows a 16-port switch) that may be used as the switching hardware engine of a basic switching unit according to an embodiment of the present invention. However, commercially available ATM switches also may operate as the switching engine of the basic switching unit according to other embodiments of the present invention. The main functional components of switching hardware 12b include a

switch core, a microcontroller complex, and a transceiver subassembly. Generally, the switch core performs the layer 2 switching, the microcontroller complex provides the system control for the ATM switch, and the transceiver subassembly provides for the interface and basic transmission and reception of signals from the physical layer. In the present example, the switch core is based on the MMC Networks ATMS 2000 ATM Switch Chip Set which includes White chip 200, Grey chip 202, MBUF chips 204, Port Interface Device (PIF) chips 206, and common data memory 208. The switch core also may optionally include VC Activity Detector 210, and Early Packet Discard function 212. Packet counters also are included but not shown. White chip 200 provides configuration control and status. In addition to communicating with White chip 200 for status and control, Grey chip 202 is responsible for direct addressing and data transfer with the switch tables. MBUF chips 204 are responsible for movement of cell traffic between PIF chips 206 and the common data memory 208. Common data memory 208 is used as cell buffering within the switch. PIF chips 206 manage transfer of data between the MBUF chips to and from the switch port hardware. VC Activity Detector 210 which includes a memory element provides information on every active virtual channel. Early Packet Discard 212 provides the ability to discard certain ATM cells as needed. Packet counters provide the switch with the ability to count all packets passing all input and output ports. Buses 214, 215, 216, 217, and 218 provide the interface between the various components of the switch. The microcontroller complex includes a central processing unit (CPU) 230, dynamic random access memory (DRAM) 232, read only memory (ROM) 234, flash memory 236, DRAM controller 238, Dual Universal Asynchronous Receiver-Transmitter (DUART) ports 240 and 242, and external timer 244. CPU 230 acts as the microcontroller. ROM 234 acts as the local boot ROM and includes the entire switch code image, basic low-level operation system functionality, and diagnostics. DRAM 232 provides conventional random access memory functions, and DRAM controller 238 (which may be implemented by a field programmable gate array (FPGA) device or the like) provides refresh control for DRAM 232. Flash memory 236 is accessible by the microcontroller for hardware revision control, serial number identification, and various control codes for manufacturability and tracking. DUART Ports 240 and 242 are provided as interfaces to communications resources for diagnostic, monitoring, and other purposes. External timer 244 interrupts CPU 230 as required. Transceiver subassembly includes physical interface devices 246, located between PIF chips 206 and physical transceivers

(not shown). Interface devices 246 perform processing of the data stream, and implement the ATM physical layer. Of course, the components of the switch may be on a printed circuit board that may reside on a rack for mounting or for setting on a desktop, depending on the chassis that may be used.

Without in any way limiting the scope of the invention, Table 2 provides a list of commercially available components which are useful in operation of the switching engine, according to the above specific embodiment. It will be apparent to those of skill in the art that the components listed in Table 2 are merely representative of those which may be used in association with the inventions herein and are provided for the purpose of facilitating assembly of a device in accordance with a particular embodiment of the invention. A wide variety of components or available switches readily known to those of skill in the art could readily be substituted or functionality could be combined or separated.

Table 2: Switch Components

SWITCH CORE	
Core chip set	MMC Networks ATMS 2000 ATM Switch Chip Set (White chip, Grey chip, MBUF chips, PIF chips)
Common data memory	standard memory modules
Packet counters	standard counters
MICROCONTROLLER COMPLEX	
CPU	Intel 960CA/CF/HX
DRAM	standard DRAM modules
ROM	standard ROM
Flash memory	standard flash memory
DRAM controller	standard FPGA, ASIC, etc.
DUART	16552 DUART
External timer	standard timer
TRANSCEIVER SUBASSEMBLY	
Physical interface	PMC-Sierra PM5346

III. System Software Functionality

As generally described above in accordance with the specific embodiment, IFMP is a protocol for instructing an adjacent node to attached a layer 2 "label" to a

system node that is a basic switching unit performs step 280. Other system nodes (e.g., switch gateway unit or host) operate as shown in Fig. 5b but do not perform step 280 since the result of step 278 is no for a switch gateway unit or a host (as these types of system nodes have no downstream link).

Fig. 6a is a diagram generally illustrating the steps involved in labelling a flow in the upstream link of a system node, such as shown by label flow step 274 of Fig. 5b. For a system node that is a switch gateway unit or a host, the system node labels a flow as shown in steps 290, 292, 300 and 276 of Fig. 6a. When the label flow step begins (step 290), the system node selects a free label x on the upstream link in step 292. The system node then sends an IFMP redirect message on the upstream link in step 300 (as indicated by dotted line 293). The system node then forwards the packet in step 276. For a system node that is a basic switching unit, labelling a flow is also illustrated by steps 294, 296, and 298. When the label flow step begins (step 290), the basic switching unit selects a free label x on the upstream link in step 292. The switch controller of basic switching unit then selects a temporary label x' on the control port of the switch controller in step 294. At step 296, the switch controller then sends to the hardware switching engine a GSMP message to map label x on the upstream link to label x' on the control port. The switch controller then waits in step 298 until a GSMP acknowledge message is received from the hardware switching engine that indicates that the mapping is successful. Upon receiving acknowledgement, the basic switching unit sends an IFMP redirect message on the upstream link in step 300. After step 300, the system node returns to step 176 as shown in Fig. 5b.

Fig. 6b is a diagram generally illustrating the steps involved in switching a flow in a basic switching unit, such as shown by switch flow step 280 of Fig. 5b. As mentioned above, only system nodes that are basic switching units may perform the switch flow step. When the switch flow procedure starts in step 310, the switch controller in the basic switching unit sends at step 312 a GSMP message to map label x on the upstream link to the label y on the downstream link. Label y is the label which the node downstream to the basic switching unit has assigned to the flow. Of course, this downstream node has labelled the flow in the manner specified by Figs. 5b and 6a, with the free label y being selected in step 292. After step 312, the switch controller in the basic switching unit waits in step 314 for a GSMP acknowledge message from a hardware switching engine in basic switching unit to indicate that the mapping is successful. The

This Page Blank (uspto)

downstream and switch controller 12a of the upstream basic switching unit 12 in step 396 determines that the flow label is for a multiport interface, then the upstream basic switching unit 12 switches the flow for the designated multiport interface in step 280. After accordingly switching the flow using the appropriate label determined downstream (for a particular port if the flow label is not for a multiport interface, or for a particular sub-port if the flow label is for a multiport interface) in step 280, the basic switching unit 12 at step 276 forwards the packet downstream.

When a packet is forwarded or switched from an upstream node and received at the downstream node, the downstream node proceeds to forward the packet traffic on the chosen label from all sub-ports of the multiport interface.

As seen in the above description for a specific embodiment, inverse multiplexing across sub-ports of a multiport interface may be accomplished on a flow-by-flow basis with the net result being that traffic is distributed fairly evenly across the sub-ports.

In accordance with another specific embodiment, inverse multiplexing across sub-ports of a multiport interface may be accomplished on a flow-by-flow basis in another manner with the net result being that traffic is distributed across the sub-ports such that the flow is not balanced across all the specified sub-ports but instead the current sub-port is desired to be fully loaded with flows before adding a flow to the next sub-port. In particular, the present specific embodiment is achieved by using weighted multiport interfaces, which are configured in a similar manner as multiport interfaces. It is noted that both the upstream and downstream basic switching units are each configured for the same weighted multiport interfaces. Configuring a weighted multiport interface may be achieved in the specific embodiment by a command to define the weighted multiport interface. Specifically, the switch controller of the upstream basic switching unit (Fig. 1) defines a weighted multiport interface by a command (e.g., *define wmpif 8 9 10 12*) which designates, in this example, ports 8, 9, 10 and 12 to be sub-ports of the weighted multiport interface. Similarly, managing the weighted multiport interface may be achieved with other commands to show the weighted multiport interface, and to delete the weighted multiport interface. In this example, when sub-port 8 of the weighted multiport interface is running at full line rate, then sub-port 9 will be opened for queuing of flows. When sub-port 9 is running at the full line rate, then sub-port 10 will be opened for queuing of flows, and so forth.

This Page Blank (uspto)

According to yet another specific embodiment, the use of virtual path interfaces (VPIs) may be combined with multiport interfaces to provide extremely high bandwidth tunneled across a wide area network (WAN). More specifically, this may be accomplished by configuring a virtual path interface that refers to a multiport interface as the carrier of the virtual path provided by the WAN. It is noted that both the upstream and downstream basic switching units are configured for the same virtual path interface referring to the same multiport interface. In accordance with the specific embodiment of the present invention, a multiport interface is first created and the basic switching unit configured, in a similar manner as described for Fig. 7a. As an example, a multiport interface is created by the command (e.g., *define mpif 8 9 10 11*) which designates, in this example, ports 8, 9, 10 and 11 to be sub-ports of multiport interface numbered as port 8. After receiving the command to define the multiport interface, the switching engine may return an acknowledgment message indicating the successful definition of the multiport interface (e.g., *multiport interface 8 successfully defined*). Then, a virtual path interface is configured in this specific embodiment by a command to define the virtual path interface. Specifically, the switch controller of the upstream basic switching unit (Fig. 1) defines a virtual path interface 17 by a command (e.g., *define vpif 8 5 17*) which designates, in this example, a virtual path interface 5 on the ports 8, 9, 10 and 11 (sub-ports of the multiport interface numbered as port 8). The above commands combine the virtual path interface 5 on ports 8-11 into a virtual path interface numbered 17. In the present specific embodiment, the same VPI number should be used on all sub-ports of the created multiport interface to carry traffic across the WAN. Managing the virtual path interfaces may be achieved with other commands to show the virtual interface, and to delete the virtual path interface. After receiving the command to define the virtual path interface combined with the multiport interface, the switching engine may return an acknowledgment message indicating the successful definition of the virtual path interface (e.g., *virtual path interface 17 on port 8 vpi 5 successfully defined*). Configuration of a virtual path interface combined with a multiport interface requires that the switching engine and the switch controller re-initialize (e.g., by rebooting) their communication in order to exchange the new list of available interfaces.

In accordance with a specific embodiment of the present invention, the source code of the system software (© Copyright, Unpublished Work, Ipsilon Networks,

This Page Blank (uspto)

Inc., All Rights Reserved) for use on the switch controller of a basic switching unit is included as Appendix I. In particular, Appendix I includes the system software for configuration and operation of multiport interfaces, flow characterization and direction on sub-ports, interfacing with IFMP and GSMP protocols, routing and forwarding, device drivers, operating system interfaces, as well as drivers and modules.

IV. Conclusion

The inventions claimed herein provide an improved method and apparatus for transmitting packets over a network by multiplexing IP switched flows over a multiport interface between basic switching units. It is to be understood that the above description is intended to be illustrative and not restrictive. Many embodiments will be apparent to those of skill in the art upon reviewing the above description. By way of example the inventions herein have been illustrated primarily with regard to transmission of IP packets capable of carrying voice, video, image, facsimile, and data signals, but they are not so limited. By way of further example, the invention has been illustrated in conjunction with specific components and operating speeds, but the invention is not so limited. The scope of the inventions should, therefore, be determined not with reference to the above description, but should instead be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled, by one of ordinary skill in the art.

This Page Blank (uspto)

WHAT IS CLAIMED IS:

1. A method for transmitting packets over a multiport interface between an upstream node and a downstream node in a network, said downstream node being downstream from said upstream node, said method comprising the steps of:

establishing a multiport interface comprising a plurality of sub-ports between said upstream node and said downstream node;

receiving a packet at said downstream node;

performing a flow classification at said downstream node on said packet to determine whether said packet belongs to a specified flow that should be redirected in the upstream node to said multiport interface;

selecting a free label for one of said plurality of sub-ports at said downstream node;

informing said upstream node that future packets belonging to said specified flow should be sent with said selected free label attached.

2. The method of claim 1 wherein said upstream and downstream nodes use ATM.

3. The method of claim 2 wherein said free label comprises a VPI/VCI.

4. The method of claim 1 wherein said network comprises a local area computer network.

5. The method of claim 1 wherein said network comprises a wide area network (WAN).

6. The method of claim 1 further comprising the step of:

determining said one of said plurality of sub-ports for which said free label is selected.

7. The method of claim 6 wherein a lowest numbered sub-port of said plurality of sub-ports is selected as said one of said plurality of sub-ports when all of said plurality

This Page Blank (uspto)

of sub-ports are running at equal outgoing cell rates.

8. The method of claim 6 wherein said one of said plurality of sub-ports has the fewest outgoing cell rate of all of said plurality of sub-ports.

9. The method of claim 6 wherein said one of said plurality of sub-ports has the shortest queue of waiting traffic at a given priority level among all of said plurality of sub-ports.

10. The method of claim 6 wherein said determining step includes ensuring that said one of said plurality of sub-ports is not a sub-port experiencing failure.

11. The method of claim 6 wherein said informing step is performed by IFMP software that enables communication between said upstream and downstream node, and said determining step uses GSMP software.

12. The method of claim 6 wherein said one of said plurality of sub-ports is not yet fully loaded with flows and each of the remaining of said plurality of sub-ports either is fully loaded with flows or is loaded with no flows.

13. The method of claim 6 further comprising the step of:

configuring a virtual path interface that refers to said multiport interface, and wherein each of said plurality of sub-ports uses said virtual path interface.

14. A computer program product that enables dynamic shifting between routing and switching in a network having an upstream node and a downstream node downstream from said upstream node, said computer program product comprising:

computer-readable code that establishes a multiport interface comprising a plurality of sub-ports between said upstream node and said downstream node;

computer-readable code that performs a flow classification on a packet at said downstream node to determine whether said packet belongs to a specified flow that should be redirected in said upstream node to said multiport interface;

computer-readable code that selects a free label for one of said plurality of

This Page Blank (uspto)

sub-ports at said downstream node;

computer-readable code that informs said upstream node that future packets belonging to said specified flow should be sent with said selected first free label attached; and

a tangible medium that stores the computer-readable codes.

15. The computer program product of claim 14, wherein said tangible media comprises a hard disk on a computer.

16. The computer program product of claim 14, wherein said tangible media is selected from a group consisting of CD-ROM, tape, floppy disk, and the like.

17. The computer program product of claim 14 wherein said computer-readable codes are installed on a computer attached to a switching hardware engine.

18. The computer program product of claim 17 wherein said computer attached to a switching hardware engine is an IP switched router.

19. The computer program product of claim 18 wherein said switching hardware engine utilizes asynchronous transfer mode (ATM) switching technology.

20. The computer program product of claim 19 wherein said flow classification uses VPI/VCI as labels.

21. The computer program product of claim 19 wherein said switching hardware engine utilizes a switching technology selected from a group consisting of FDDI, Ethernet, Fast Ethernet, Gigabit Ethernet, frame relay, and fast packet switching.

This Page Blank (uspto)

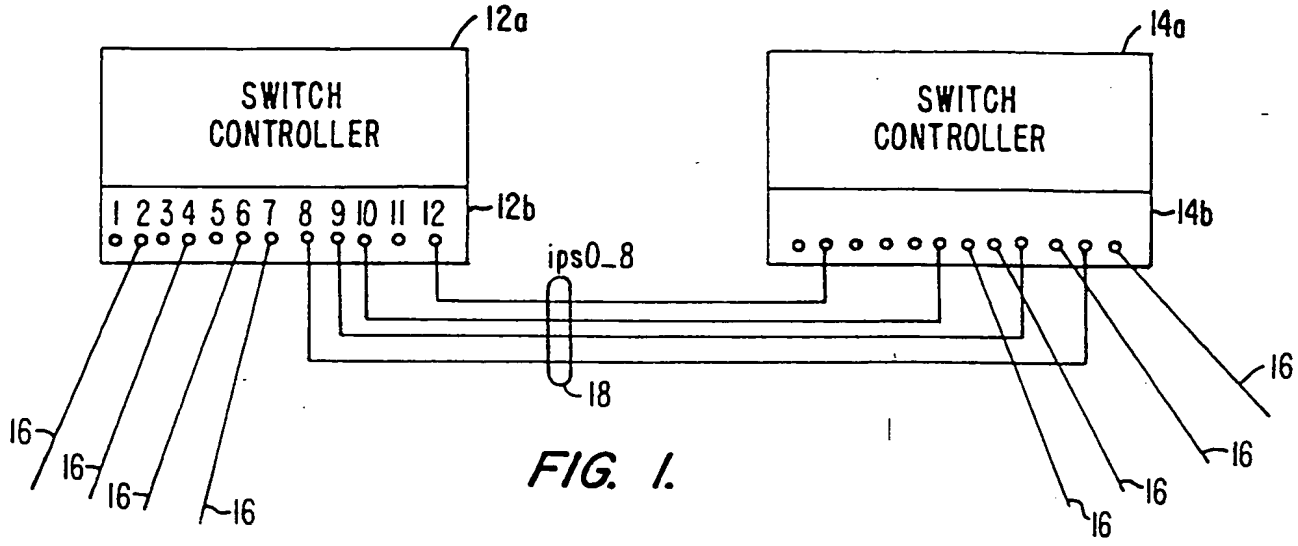

12a
SWITCH CONTROLLER
1 2 3 4 5 6 7 8 9 10 11 12
12b
ipsO_8
18
14a
SWITCH CONTROLLER
14b
16
16
16
16
16
16
16
16

FIG. 1.

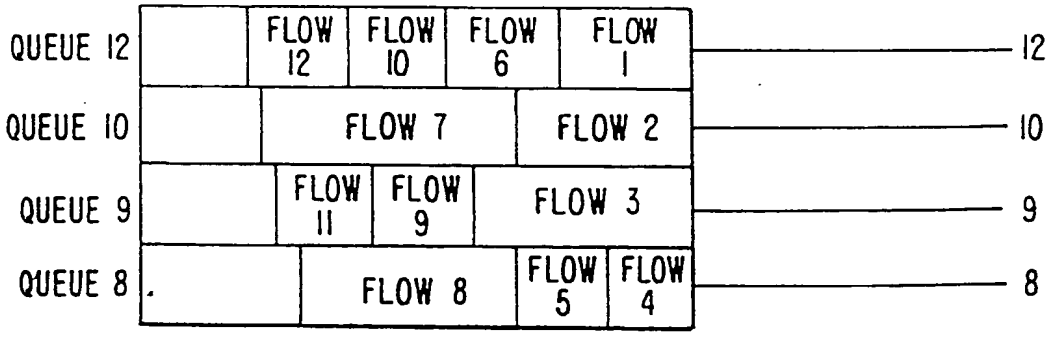

QUEUE 12
FLOW 12
FLOW 10
FLOW 6
FLOW 1
12
QUEUE 10
FLOW 7
FLOW 2
10
QUEUE 9
FLOW 11
FLOW 9
FLOW 3
9
QUEUE 8
FLOW 8
FLOW 5
FLOW 4
8

FIG. 2.

This Page Blank (uspto)

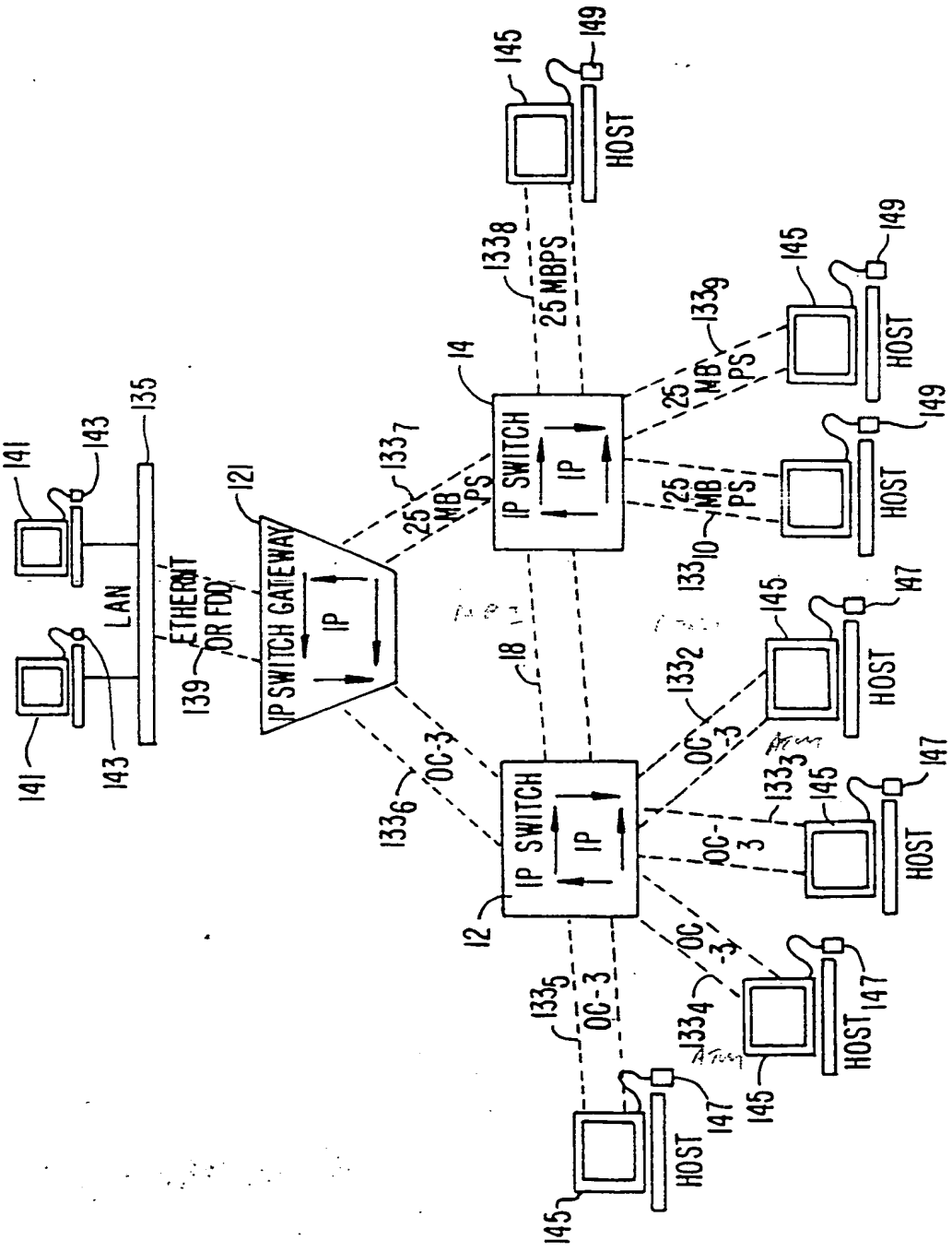

IP SWITCH GATEWAY
IP
IP SWITCH
IP
IP SWITCH
IP
LAN
ETHERNET OR FDDI
HOST
25 MBPS
25 MB PS
OC-3
OC-1 3
121
135
139
141
143
145
147
149
12
14
18
133_2
133_3
133_4
133_5
133_6
133_7
133_8
133_9
133_{10}

FIG. 3.

This Page Blank (uspto)

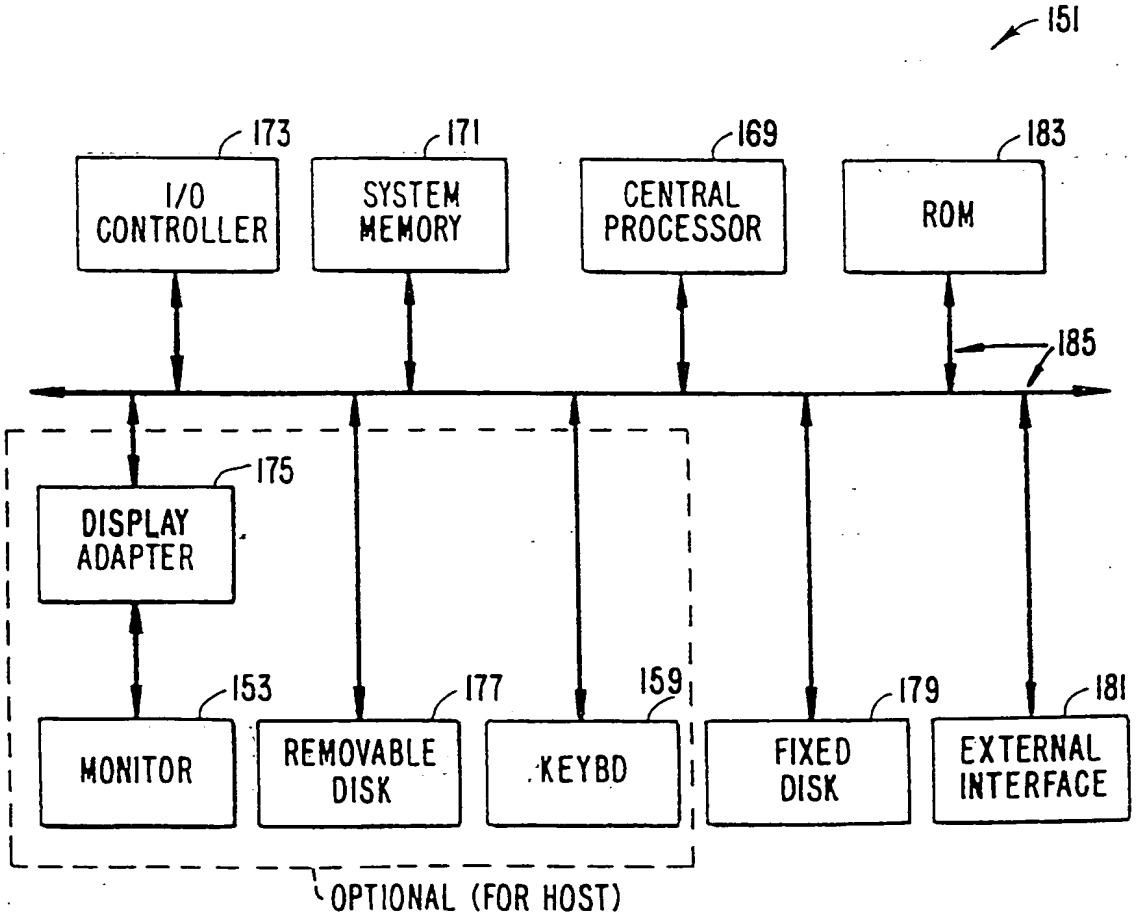

151
173
I/O CONTROLLER
171
SYSTEM MEMORY
169
CENTRAL PROCESSOR
183
ROM
185
175
DISPLAY ADAPTER
153
MONITOR
177
REMOVABLE DISK
159
KEYBD
179
FIXED DISK
181
EXTERNAL INTERFACE
OPTIONAL (FOR HOST)

FIG. 4A.

This Page Blank (uspto)

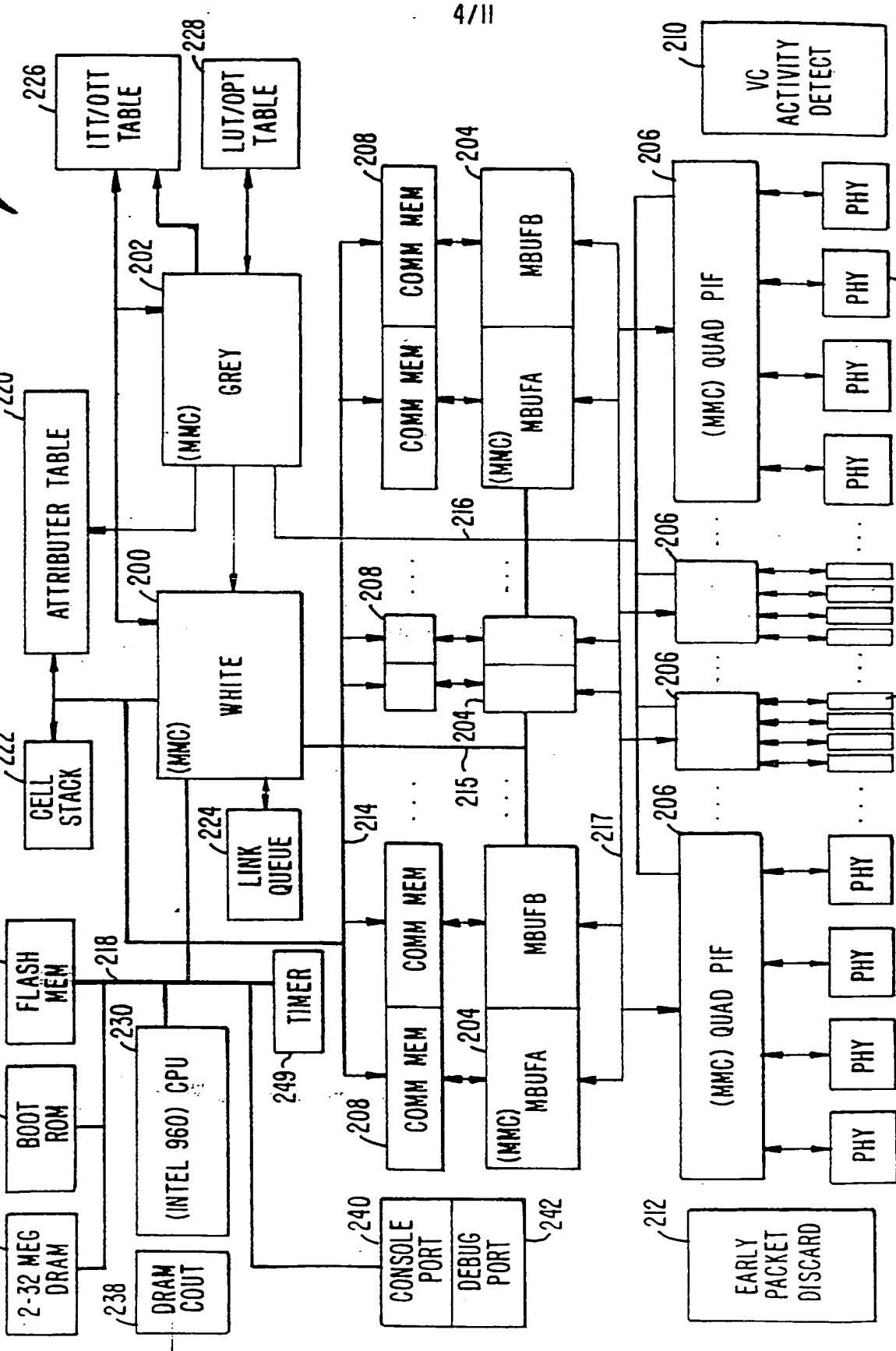

126
ITT/OTT TABLE
226
LUT/OPT TABLE
228
VC ACTIVITY DETECT
210
ATTRIBUTER TABLE
220
(MMC) GREY
202
COMM MEM
208
MBUFB
204
(MMC) QUAD PIF
206
PHY
246
(MMC) WHITE
200
MBUFA
CELL STACK
222
LINK QUEUE
224
216
214
215
217
FLASH MEM
236
218
TIMER
249
BOOT ROM
234
(INTEL 960) CPU
230
2-32 MEG DRAM
232
DRAM COUT
238
CONSOLE PORT
240
DEBUG PORT
242
EARLY PACKET DISCARD
212

FIG. 4B.

This Page Blank (uspto)

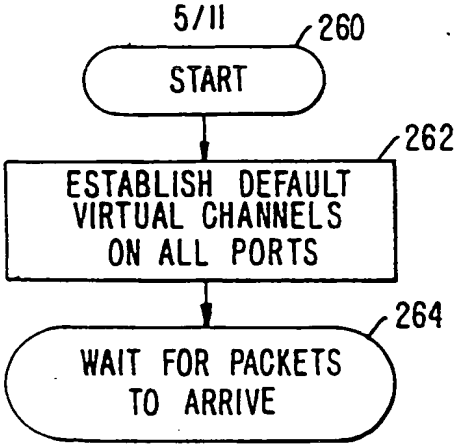

260
START
262
ESTABLISH DEFAULT VIRTUAL CHANNELS ON ALL PORTS
264
WAIT FOR PACKETS TO ARRIVE

FIG. 5A.

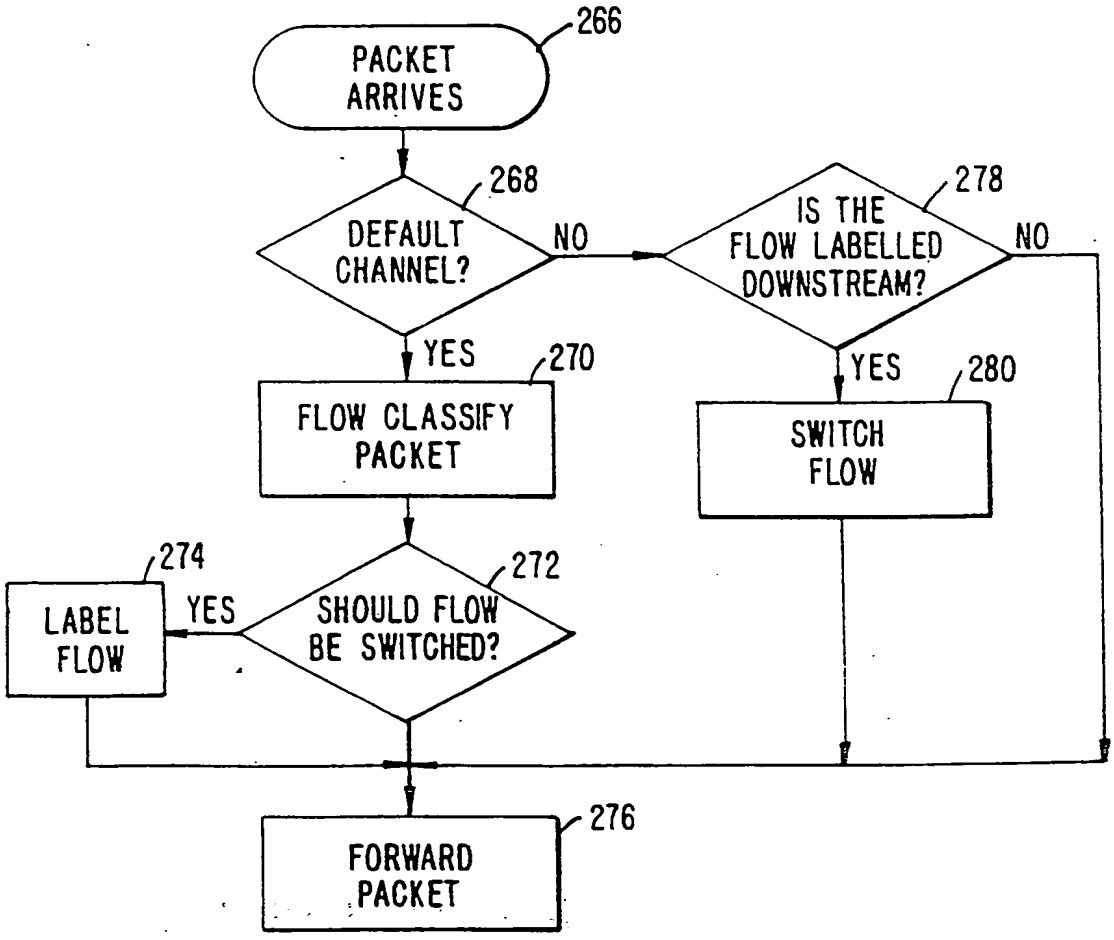

266
PACKET ARRIVES
268
DEFAULT CHANNEL?
NO
278
IS THE FLOW LABELLED DOWNSTREAM?
NO
YES
270
FLOW CLASSIFY PACKET
YES
280
SWITCH FLOW
274
LABEL FLOW
YES
272
SHOULD FLOW BE SWITCHED?
276
FORWARD PACKET

FIG. 5B.

This Page Blank (uspto)

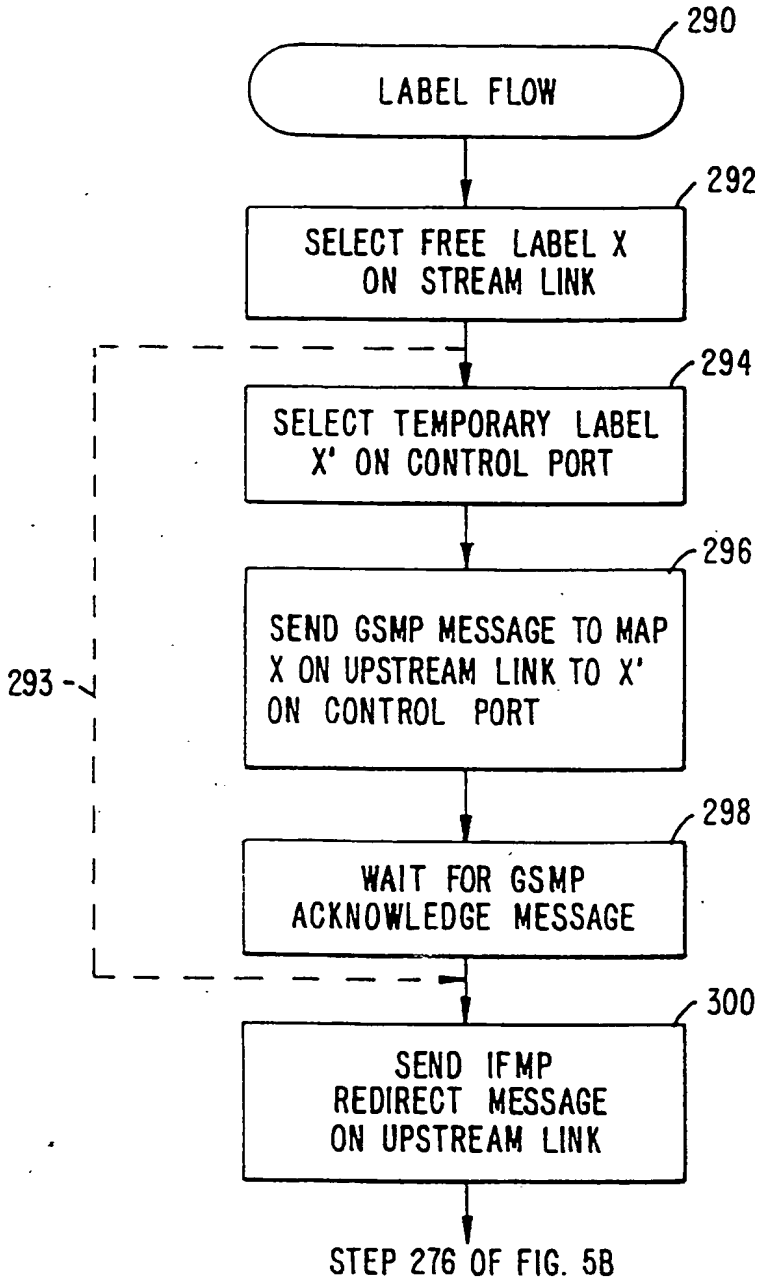

290
LABEL FLOW
292
SELECT FREE LABEL X ON STREAM LINK
294
SELECT TEMPORARY LABEL X' ON CONTROL PORT
296
SEND GSMP MESSAGE TO MAP X ON UPSTREAM LINK TO X' ON CONTROL PORT
293
298
WAIT FOR GSMP ACKNOWLEDGE MESSAGE
300
SEND IFMP REDIRECT MESSAGE ON UPSTREAM LINK
STEP 276 OF FIG. 5B

FIG. 6A.

This Page Blank (uspto)

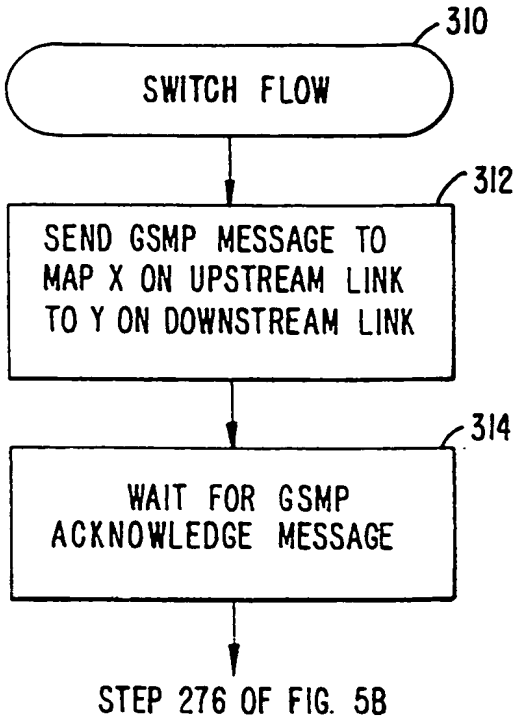

310
SWITCH FLOW
312
SEND GSMP MESSAGE TO MAP X ON UPSTREAM LINK TO Y ON DOWNSTREAM LINK
314
WAIT FOR GSMP ACKNOWLEDGE MESSAGE
STEP 276 OF FIG. 5B

FIG. 6B.

This Page Blank (uspto)

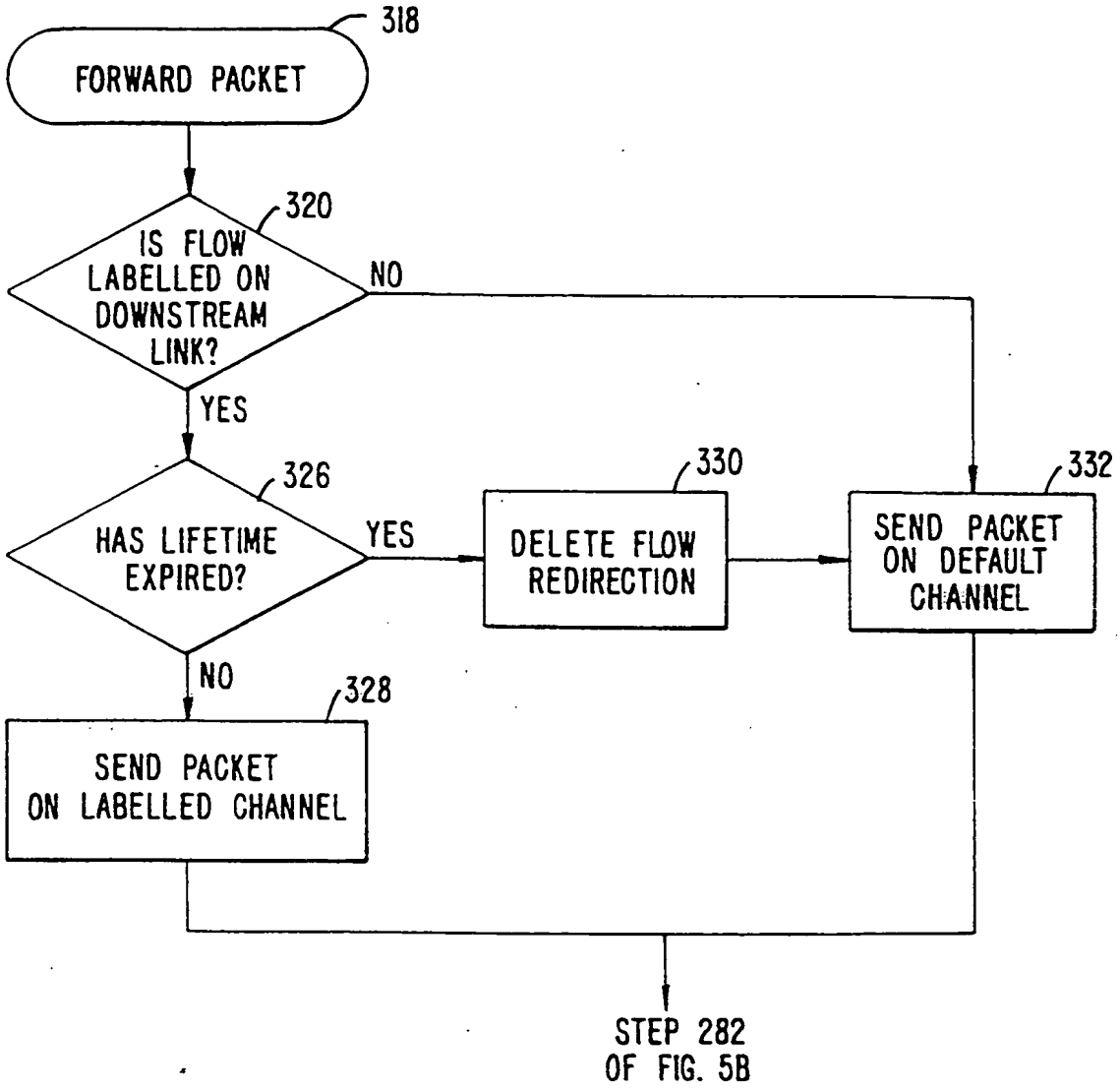

318
FORWARD PACKET
320
IS FLOW LABELLED ON DOWNSTREAM LINK?
NO
YES
326
HAS LIFETIME EXPIRED?
YES
330
DELETE FLOW REDIRECTION
332
SEND PACKET ON DEFAULT CHANNEL
NO
328
SEND PACKET ON LABELLED CHANNEL
STEP 282 OF FIG. 5B

FIG. 6C.

This Page Blank (uspto)

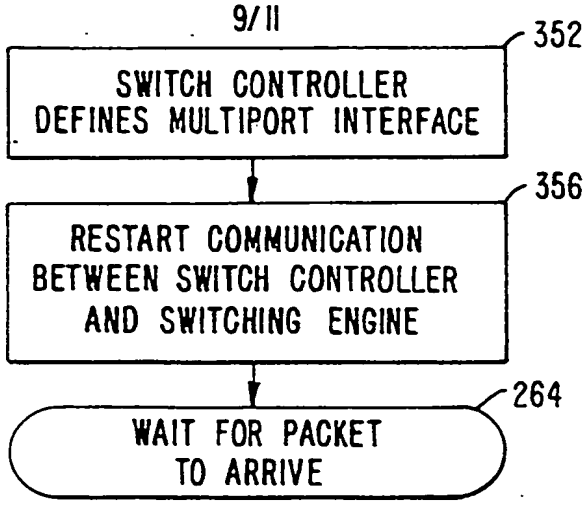

352
SWITCH CONTROLLER DEFINES MULTIPORT INTERFACE
356
RESTART COMMUNICATION BETWEEN SWITCH CONTROLLER AND SWITCHING ENGINE
264
WAIT FOR PACKET TO ARRIVE

FIG. 7A.

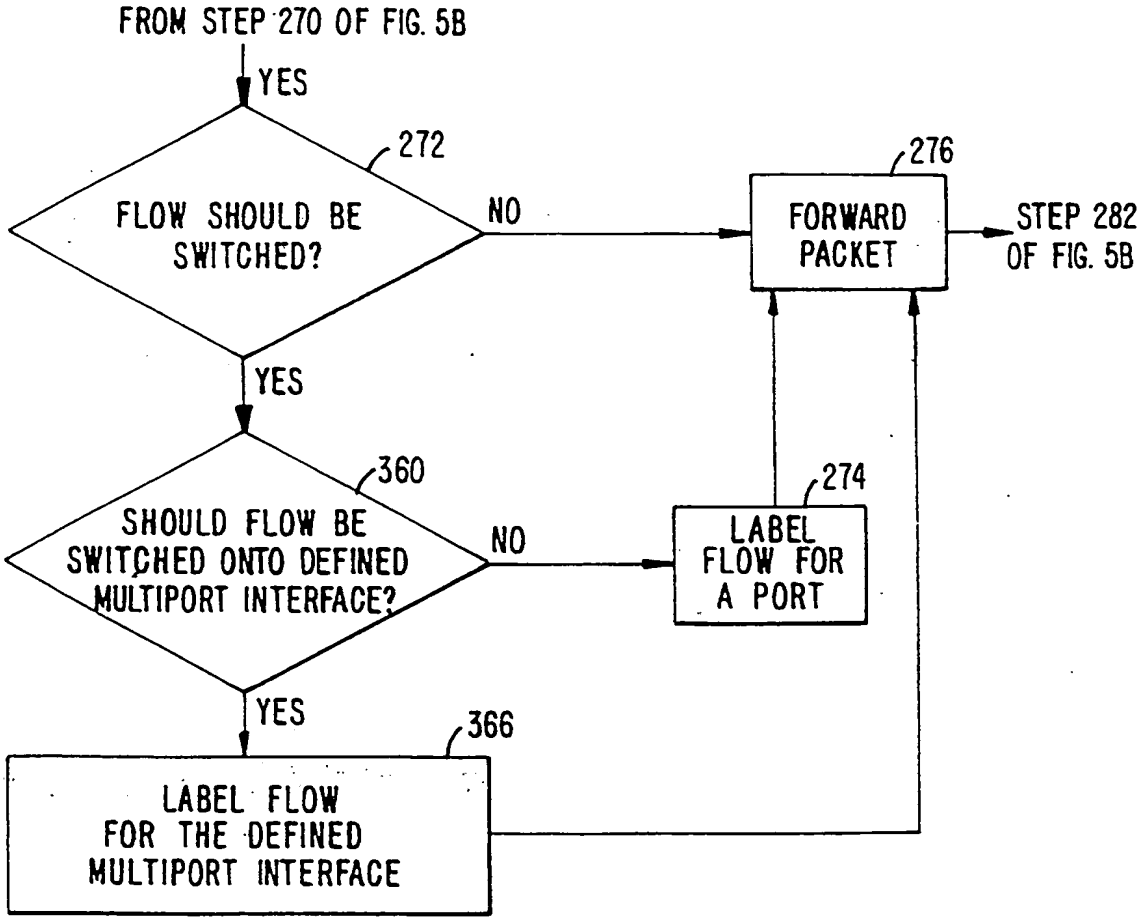

FROM STEP 270 OF FIG. 5B
YES
272
FLOW SHOULD BE SWITCHED?
NO
276
FORWARD PACKET
STEP 282 OF FIG. 5B
YES
360
SHOULD FLOW BE SWITCHED ONTO DEFINED MULTIPORT INTERFACE?
NO
274
LABEL FLOW FOR A PORT
YES
366
LABEL FLOW FOR THE DEFINED MULTIPORT INTERFACE

FIG. 7B.

This Page Blank (uspto)

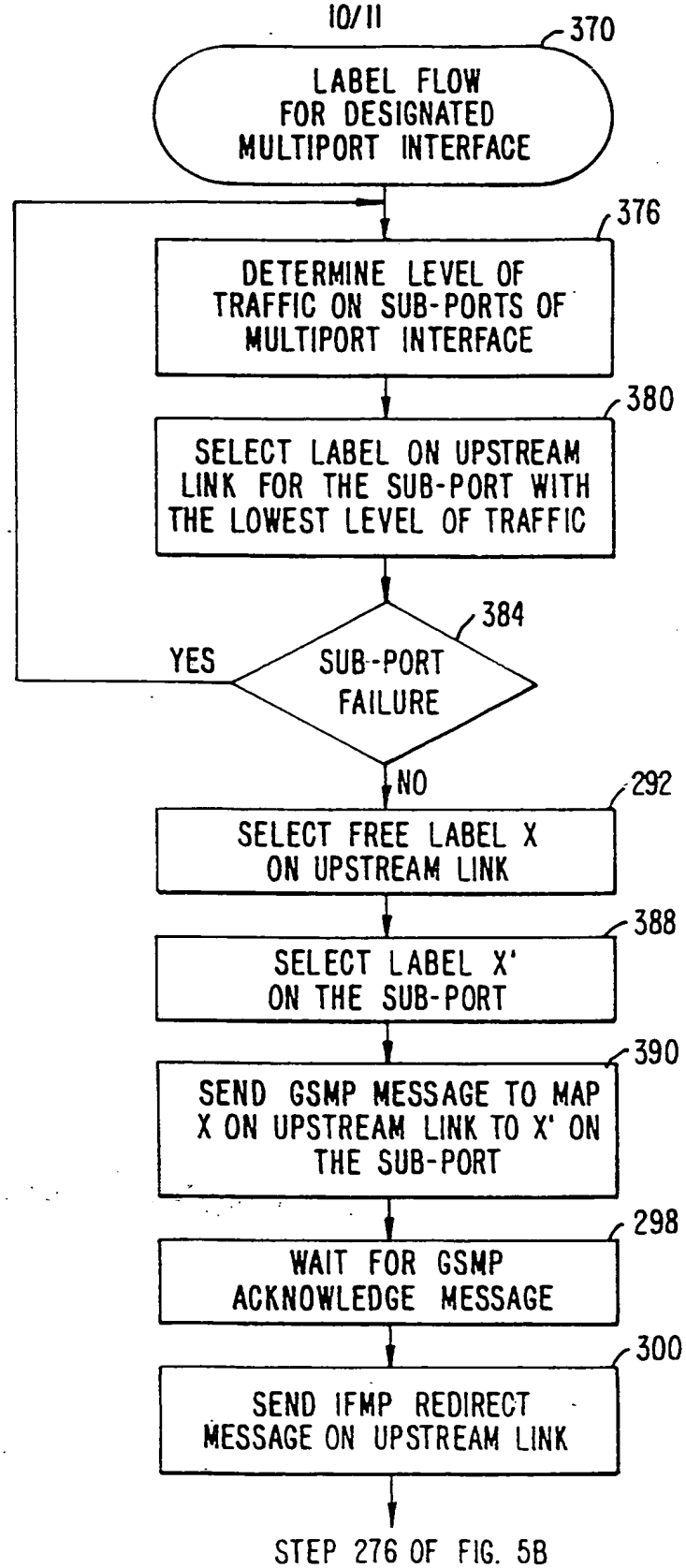

370
LABEL FLOW FOR DESIGNATED MULTIPORT INTERFACE
376
DETERMINE LEVEL OF TRAFFIC ON SUB-PORTS OF MULTIPORT INTERFACE
380
SELECT LABEL ON UPSTREAM LINK FOR THE SUB-PORT WITH THE LOWEST LEVEL OF TRAFFIC
384
SUB-PORT FAILURE
YES
NO
292
SELECT FREE LABEL X ON UPSTREAM LINK
388
SELECT LABEL X' ON THE SUB-PORT
390
SEND GSMP MESSAGE TO MAP X ON UPSTREAM LINK TO X' ON THE SUB-PORT
298
WAIT FOR GSMP ACKNOWLEDGE MESSAGE
300
SEND IFMP REDIRECT MESSAGE ON UPSTREAM LINK
STEP 276 OF FIG. 5B

FIG. 7C.

This Page Blank (uspto)

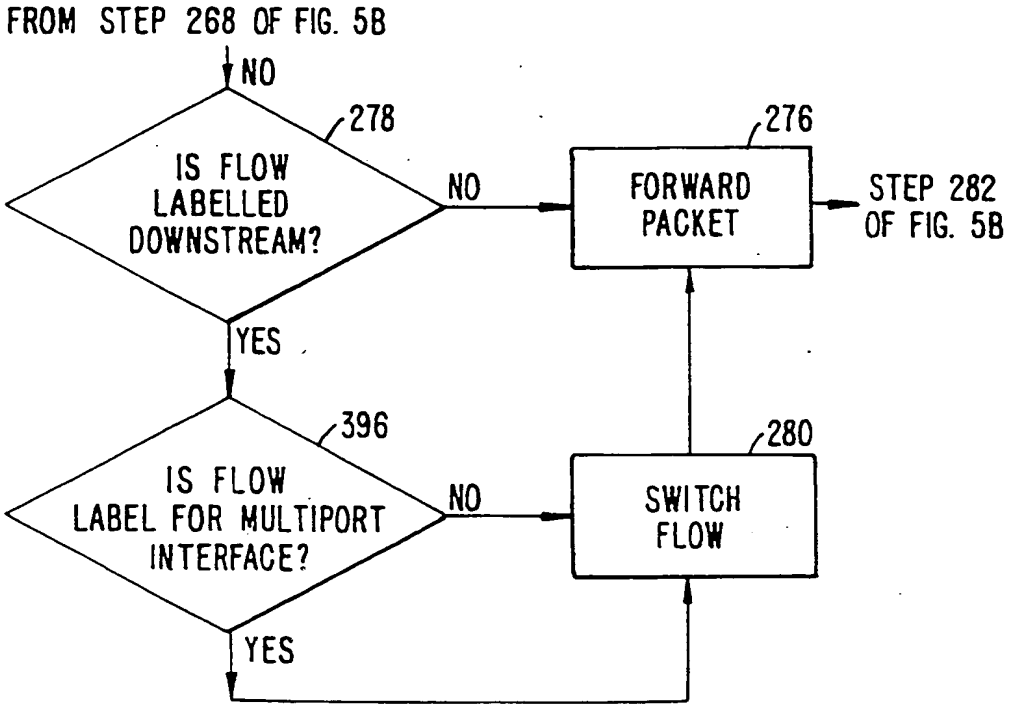

FROM STEP 268 OF FIG. 5B
NO
278
IS FLOW LABELLED DOWNSTREAM?
NO
276
FORWARD PACKET
STEP 282 OF FIG. 5B
YES
396
IS FLOW LABEL FOR MULTIPORT INTERFACE?
NO
280
SWITCH FLOW
YES

FIG. 7D.

This Page Blank (uspto)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/23535

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04Q11/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 28505 A (IPSILON NETWORKS INC) 7 August 1997 cited in the application see page 3, line 26 - page 5, line 20 see page 10, line 16 - line 20 see page 12, line 10 - line 32 see page 18, line 2 - page 20, line 24	1-6,9, 13-21
A	WO 95 27384 A (ARNOLD JOHN SPENCER ;PLESSEY TELECOMM (GB)) 12 October 1995 see page 3, line 22 - page 8, line 9	1-21
A	US 5 568 479 A (WATANABE YOSHIHIRO ET AL) 22 October 1996 see column 12, line 39 - column 13, line 14	1-21

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 March 1999	17/03/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gregori, S

1

This Page Blank (uspto)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/23535

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9728505 A	07-08-1997	AU 1851497 A	22-08-1997
		CA 2243359 A	07-08-1997
WO 9527384 A	12-10-1995	AU 699582 B	10-12-1998
		AU 2077395 A	23-10-1995
		CN 1145157 A	12-03-1997
		DE 69505241 D	12-11-1998
		EP 0753237 A	15-01-1997
		ES 2123240 T	01-01-1999
		GB 2288297 A,B	11-10-1995
		JP 9511111 T	04-11-1997
		NO 964079 A	27-09-1996
		ZA 9502588 A	21-12-1995
US 5568479 A	22-10-1996	JP 7015462 A	17-01-1995

This Page Blank (uspto)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)